RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/697,340

DATE: 07/18/2002
TIME: 13:00:19

Input Set : **A:\501d4.app**
Output Set: **N:\CRF3\07182002\I697340.raw**

```
 4 <110> APPLICANT: Brunkow, Mary E.
 5        Jeffery, Eric W.
 6        Hjerrild, Kathryn A.
 7        Ramsdell, Fred
11 <120> TITLE OF INVENTION: IDENTIFICATION OF THE GENE CAUSING THE
12        MOUSE SCURFY PHENOTYPE AND ITS HUMAN ORTHOLOG
15 <130> FILE REFERENCE: 240083.501D4
17 <140> CURRENT APPLICATION NUMBER: US 09/697,340
18 <141> CURRENT FILING DATE: 2000-10-24
20 <160> NUMBER OF SEQ ID NOS: 14
22 <170> SOFTWARE: FastSEQ for Windows Version 3.0
24 <210> SEQ ID NO: 1
25 <211> LENGTH: 2160
26 <212> TYPE: DNA
27 <213> ORGANISM: Mus musculus
29 <400> SEQUENCE: 1
30  gctgatcccc ctctagcagt ccacttcacc aaggtgagcg agtgtccctg ctctcccca     60
31  ccagacacag ctctgctggc gaaagtggca gagaggtatt gagggtgggt gtcaggagcc    120
32  caccagtaca gctggaaaca cccagccact ccagctcccg gcaacttctc ctgactctgc    180
33  cttcagacga gacttggaag acagtcacat ctcagcagct cctctgccgt tatccagcct    240
34  gcctctgaca agaacccaat gcccaaccct aggccagcca agcctatggc tccttccttg    300
35  gcccttggcc catccccagg agtcttgcca agctggaaga ctgcacccaa gggctcagaa    360
36  cttctaggga ccaggggctc tggggggaccc ttccaaggtc gggacctgcg aagtggggcc   420
37  cacacctctt cttccttgaa ccccctgcca ccatcccagc tgcagctgcc tacagtgccc    480
38  ctagtcatgg tggcaccgtc tggggcccga ctaggtccct cacccacct acaggccctt     540
39  ctccaggaca gaccacactt catgcatcag ctctccactg tggatgccca tgcccagacc    600
40  cctgtgctcc aagtgcgtcc actggacaac ccagccatga tcagcctccc accaccttct    660
41  gctgccactg gggtcttctc cctcaaggcc cggcctggcc tgccacctgg gatcaatgtg    720
42  gccagtctgg aatgggtgtc cagggagcca gctctactct gcaccttccc acgctcgggt    780
43  acacccagga aagacagcaa ccttttggct gcaccccaag gatcctaccc actgctggca    840
44  aatggagtct gcaagtggcc tggttgtgag aaggtcttcg aggagccaga agagtttctc    900
45  aagcactgcc aagcagatca tctcctggat gagaaaggca aggcccagtg cctcctccag    960
46  agagaagtgg tgcagtctct ggagcagcag ctggagctgg aaaaggagaa gctgggagct   1020
47  atgcaggccc acctggctgg gaagatggcg ctggccaagg ctccatctgt ggcctcaatg   1080
48  gacaagagct cttgctgcat cgtagccacc agtactcagg gcagtgtgct cccggcctgg   1140
49  tctgctcctc gggaggctcc agacggcggc ctgtttgcag tgcggaggca cctctgggga   1200
50  agccatggca atagttcctt cccagagttc ttccacaaca tggactactt caagtaccac   1260
51  aatatgcgac ccccttttcac ctatgccacc cttatccgat gggccatcct ggaagccccg   1320
52  gagaggcaga ggacactcaa tgaaatctac cattggttta ctcgcatgtt cgcctacttc   1380
53  agaaaccacc ccgccacctg gaagaatgcc atccgccaca acctgagcct gcacaagtgc   1440
54  tttgtgcgag tggagagcga gaagggagca gtgtggaccg tagatgaatt tgagtttcgc   1500
55  aagaagagga gccaacgccc caacaagtgc tccaatccct gcccttgacc tcaaaaccaa   1560
56  gaaaaggtgg gcggggagg gggccaaaac catgagactg aggctgtggg ggcaaggagg   1620
```

RAW SEQUENCE LISTING   DATE: 07/18/200☐
PATENT APPLICATION: US/09/697,340   TIME: 13:00:19

         Input Set : A:\501d4.app
         Output Set: N:\CRF3\07182002\I697340.raw

```
57  caagtcctac gtgtacctat ggaaaccggg cgatgatgtg cctgctatca gggcctctgc    1680
58  tccctatcta gctgccctcc tagatcatat catctgcctt acagctgaga ggggtgccaa    1740
59  tcccagccta gcccctagtt ccaacctagc cccaagatga actttccagt caaagagccc    1800
60  tcacaaccag ctatacatat ctgccttggc cactgccaag cagaaagatg acagacacca    1860
61  tcctaatatt tactcaaccc aaaccctaaa acatgaagag cctgccttgg tacattcgtg    1920
62  aactttcaaa gttagtcatg cagtcacaca tgactgcagt cctactgact cacaccccaa    1980
63  agcactcacc cacaacatct ggaaccacgg gcactatcac acataggtgt atatacagac    2040
64  ccttacacag caacagcact ggaaccttca caattacatc cccccaaacc acacaggcat    2100
65  aactgatcat acgcagcctc aagcaatgcc caaaatacaa gtcagacaca gcttgtcaga    2160
67  <210> SEQ ID NO: 2
68  <211> LENGTH: 429
69  <212> TYPE: PRT
70  <213> ORGANISM: Mus musculus
72  <400> SEQUENCE: 2
73  Met Pro Asn Pro Arg Pro Ala Lys Pro Met Ala Pro Ser Leu Ala Leu
74   1               5                  10                  15
75  Gly Pro Ser Pro Gly Val Leu Pro Ser Trp Lys Thr Ala Pro Lys Gly
76               20                  25                  30
77  Ser Glu Leu Leu Gly Thr Arg Gly Ser Gly Gly Pro Phe Gln Gly Arg
78           35                  40                  45
79  Asp Leu Arg Ser Gly Ala His Thr Ser Ser Ser Leu Asn Pro Leu Pro
80       50                  55                  60
81  Pro Ser Gln Leu Gln Leu Pro Thr Val Pro Leu Val Met Val Ala Pro
82  65                  70                  75                  80
83  Ser Gly Ala Arg Leu Gly Pro Ser Pro His Leu Gln Ala Leu Leu Gln
84               85                  90                  95
85  Asp Arg Pro His Phe Met His Gln Leu Ser Thr Val Asp Ala His Ala
86           100                 105                 110
87  Gln Thr Pro Val Leu Gln Val Arg Pro Leu Asp Asn Pro Ala Met Ile
88           115                 120                 125
89  Ser Leu Pro Pro Pro Ser Ala Ala Thr Gly Val Phe Ser Leu Lys Ala
90       130                 135                 140
91  Arg Pro Gly Leu Pro Pro Gly Ile Asn Val Ala Ser Leu Glu Trp Val
92  145                 150                 155                 160
93  Ser Arg Glu Pro Ala Leu Leu Cys Thr Phe Pro Arg Ser Gly Thr Pro
94               165                 170                 175
95  Arg Lys Asp Ser Asn Leu Leu Ala Ala Pro Gln Gly Ser Tyr Pro Leu
96               180                 185                 190
97  Leu Ala Asn Gly Val Cys Lys Trp Pro Gly Cys Glu Lys Val Phe Glu
98           195                 200                 205
99  Glu Pro Glu Glu Phe Leu Lys His Cys Gln Ala Asp His Leu Leu Asp
100          210                 215                 220
101  Glu Lys Gly Lys Ala Gln Cys Leu Leu Gln Arg Glu Val Val Gln Ser
102  225                 230                 235                 240
103  Leu Glu Gln Gln Leu Glu Leu Glu Lys Glu Lys Leu Gly Ala Met Gln
104               245                 250                 255
105  Ala His Leu Ala Gly Lys Met Ala Leu Ala Lys Ala Pro Ser Val Ala
106               260                 265                 270
107  Ser Met Asp Lys Ser Ser Cys Cys Ile Val Ala Thr Ser Thr Gln Gly
```

RAW SEQUENCE LISTING   DATE: 07/18/2002
PATENT APPLICATION: US/09/697,340   TIME: 13:00:19

Input Set : **A:\501d4.app**
Output Set: **N:\CRF3\07182002\I697340.raw**

```
108         275              280              285
109  Ser Val Leu Pro Ala Trp Ser Ala Pro Arg Glu Ala Pro Asp Gly Gly
110      290              295     ,        300
111  Leu Phe Ala Val Arg Arg His Leu Trp Gly Ser His Gly Asn Ser Ser
112  305              310              315              320
113  Phe Pro Glu Phe Phe His Asn Met Asp Tyr Phe Lys Tyr His Asn Met
114              325              330              335
115  Arg Pro Pro Phe Thr Tyr Ala Thr Leu Ile Arg Trp Ala Ile Leu Glu
116              340              345              350
117  Ala Pro Glu Arg Gln Arg Thr Leu Asn Glu Ile Tyr His Trp Phe Thr
118          355              360              365
119  Arg Met Phe Ala Tyr Phe Arg Asn His Pro Ala Thr Trp Lys Asn Ala
120      370              375              380
121  Ile Arg His Asn Leu Ser Leu His Lys Cys Phe Val Arg Val Glu Ser
122  385              390              395              400
123  Glu Lys Gly Ala Val Trp Thr Val Asp Glu Phe Glu Phe Arg Lys Lys
124              405              410              415
125  Arg Ser Gln Arg Pro Asn Lys Cys Ser Asn Pro Cys Pro
126              420              425
128 <210> SEQ ID NO: 3
129 <211> LENGTH: 1869
130 <212> TYPE: DNA
131 <213> ORGANISM: Homo sapien
133 <400> SEQUENCE: 3
134  gcacacactc atcgaaaaaa atttggatta ttagaagaga gaggtctgcg gcttccacac     60
135  cgtacagcgt ggttttttctt ctcggtataa aagcaaagtt gtttttgata cgtgacagtt    120
136  tcccacaagc caggctgatc ctttttctgtc agtccacttc accaagcctg cccttggaca    180
137  aggacccgat gcccaacccc aggcctggca agccctcggc cccttccttg gcccttggcc    240
138  catccccagg agcctcgccc agctggaggg ctgcacccaa agcctcagac ctgctggggg    300
139  cccgggggccc aggggggaacc ttccagggcc gagatcttcg aggcggggcc catgcctcct    360
140  cttcttcctt gaaccccatg ccaccatcgc agctgcagct gcccacactg cccctagtca    420
141  tggtggcacc ctccggggca cggctgggcc ccttgcccca cttacaggca ctcctccagg    480
142  acaggccaca tttcatgcac cagctctcaa cggtggatgc ccacgcccgg accctgtgc     540
143  tgcaggtgca cccctggag agcccagcca tgatcagcct cacaccaccc accaccgcca    600
144  ctggggtctt ctccctcaag gcccggcctg gcctcccacc tgggatcaac gtggccagcc    660
145  tggaatgggt gtccagggag ccggcactgc tctgcacctt cccaaatccc agtgcaccca    720
146  ggaaggacag caccctttcg gctgtgcccc agagctccta cccactgctg gcaaatggtg    780
147  tctgcaagtg gcccggatgt gagaaggtct tcgaagagcc agaggacttc ctcaagcact    840
148  gccaggcgga ccatcttctg gatgagaagg gcagggcaca atgtctcctc cagagagaga    900
149  tggtacagtc tctggagcag cagctggtgc tggagaagga gaagctgagt gccatgcagg    960
150  cccacctggc tgggaaaatg gcactgacca aggcttcatc tgtggcatca tccgacaagg    1020
151  gctcctgctg catcgtagct gctggcagcc aaggccctgt cgtcccagcc tggtctggcc    1080
152  cccgggaggc ccctgacagc ctgtttgctg tccggaggca cctgtggggt agccatggaa    1140
153  acagcacatt cccagagttc ctccacaaca tggactactt caagttccac aacatgcgac    1200
154  cccctttcac ctacgccacg ctcatccgct gggccatcct ggaggctcca gagaagcagc    1260
155  ggacactcaa tgagatctac cactggttca cacgcatgtt tgccttcttc agaaaccatc    1320
156  ctgccacctg gaagaacgcc atccgccaca acctgagtct gcacaagtgc tttgtgcggg    1380
157  tggagagcga gaagggggct gtgtggaccg tggatgagct ggagttccgc aagaaacgga    1440
158  gccagaggcc cagcaggtgt tccaacccta cacctggccc ctgacctcaa gatcaaggaa    1500
```

Input Set : **A:\501d4.app**
Output Set: **N:\CRF3\07182002\I697340.raw**

```
159  aggaggatgg acgaacaggg gccaaactgg tgggaggcag aggtggtggg ggcagggatg    1560
160  ataggccctg gatgtgccca cagggaccaa gaagtgaggt ttccactgtc ttgcctgcca    1620
161  gggcccctgt tcccccgctg gcagccaccc cctcccccat catatccttt gccccaaggc    1680
162  tgctcagagg ggccccggtc ctggccccag cccccacctc cgccccagac acaccccca     1740
163  gtcgagccct gcagccaaac agagccttca caaccagcca cacagagcct gcctcagctg    1800
164  ctcgcacaga ttacttcagg gctggaaaag tcacacagac acacaaatg tcacaatcct      1860
165  gtccctcac                                                             1869
167  <210> SEQ ID NO: 4
168  <211> LENGTH: 431
169  <212> TYPE: PRT
170  <213> ORGANISM: Homo sapien
172  <400> SEQUENCE: 4
173  Met Pro Asn Pro Arg Pro Gly Lys Pro Ser Ala Pro Ser Leu Ala Leu
174   1               5                  10                  15
175  Gly Pro Ser Pro Gly Ala Ser Pro Ser Trp Arg Ala Ala Pro Lys Ala
176              20                  25                  30
177  Ser Asp Leu Leu Gly Ala Arg Gly Pro Gly Gly Thr Phe Gln Gly Arg
178          35                  40                  45
179  Asp Leu Arg Gly Gly Ala His Ala Ser Ser Ser Ser Leu Asn Pro Met
180      50                  55                  60
181  Pro Pro Ser Gln Leu Gln Leu Pro Thr Leu Pro Leu Val Met Val Ala
182  65                  70                  75                  80
183  Pro Ser Gly Ala Arg Leu Gly Pro Leu Pro His Leu Gln Ala Leu Leu
184              85                  90                  95
185  Gln Asp Arg Pro His Phe Met His Gln Leu Ser Thr Val Asp Ala His
186              100                 105                 110
187  Ala Arg Thr Pro Val Leu Gln Val His Pro Leu Glu Ser Pro Ala Met
188              115                 120                 125
189  Ile Ser Leu Thr Pro Pro Thr Thr Ala Thr Gly Val Phe Ser Leu Lys
190          130                 135                 140
191  Ala Arg Pro Gly Leu Pro Pro Gly Ile Asn Val Ala Ser Leu Glu Trp
192  145                 150                 155                 160
193  Val Ser Arg Glu Pro Ala Leu Leu Cys Thr Phe Pro Asn Pro Ser Ala
194              165                 170                 175
195  Pro Arg Lys Asp Ser Thr Leu Ser Ala Val Pro Gln Ser Ser Tyr Pro
196              180                 185                 190
197  Leu Leu Ala Asn Gly Val Cys Lys Trp Pro Gly Cys Glu Lys Val Phe
198              195                 200                 205
199  Glu Glu Pro Glu Asp Phe Leu Lys His Cys Gln Ala Asp His Leu Leu
200          210                 215                 220
201  Asp Glu Lys Gly Arg Ala Gln Cys Leu Leu Gln Arg Glu Met Val Gln
202  225                 230                 235                 240
203  Ser Leu Glu Gln Gln Leu Val Leu Glu Lys Glu Lys Leu Ser Ala Met
204              245                 250                 255
205  Gln Ala His Leu Ala Gly Lys Met Ala Leu Thr Lys Ala Ser Ser Val
206              260                 265                 270
207  Ala Ser Ser Asp Lys Gly Ser Cys Cys Ile Val Ala Ala Gly Ser Gln
208              275                 280                 285
209  Gly Pro Val Val Pro Ala Trp Ser Gly Pro Arg Glu Ala Pro Asp Ser
```

RAW SEQUENCE LISTING   DATE: 07/18/2002
PATENT APPLICATION: US/09/697,340   TIME: 13:00:19

Input Set : A:\501d4.app
Output Set: N:\CRF3\07182002\I697340.raw

```
210      290                 295                 300
211 Leu Phe Ala Val Arg Arg His Leu Trp Gly Ser His Gly Asn Ser Thr
212 305                 310                 315                 320
213 Phe Pro Glu Phe Leu His Asn Met Asp Tyr Phe Lys Phe His Asn Met
214                 325                 330                 335
215 Arg Pro Pro Phe Thr Tyr Ala Thr Leu Ile Arg Trp Ala Ile Leu Glu
216                 340                 345                 350
217 Ala Pro Glu Lys Gln Arg Thr Leu Asn Glu Ile Tyr His Trp Phe Thr
218         355                 360                 365
219 Arg Met Phe Ala Phe Phe Arg Asn His Pro Ala Thr Trp Lys Asn Ala
220         370                 375                 380
221 Ile Arg His Asn Leu Ser Leu His Lys Cys Phe Val Arg Val Glu Ser
222 385                 390                 395                 400
223 Glu Lys Gly Ala Val Trp Thr Val Asp Glu Leu Glu Phe Arg Lys Lys
224                 405                 410                 415
225 Arg Ser Gln Arg Pro Ser Arg Cys Ser Asn Pro Thr Pro Gly Pro
226                 420                 425                 430
```
228 <210> SEQ ID NO: 5
229 <211> LENGTH: 23
230 <212> TYPE: DNA
231 <213> ORGANISM: Artificial Sequence
233 <220> FEATURE:
234 <223> OTHER INFORMATION: Primer for generation of mouse Fkh cDNA
236 <400> SEQUENCE: 5
237 gcagatctcc tgactctgcc ttc   23
239 <210> SEQ ID NO: 6
240 <211> LENGTH: 23
241 <212> TYPE: DNA
242 <213> ORGANISM: Artificial Sequence
244 <220> FEATURE:
245 <223> OTHER INFORMATION: Primer for generation of mouse Fkh cDNA
247 <400> SEQUENCE: 6
248 gcagatctga caagctgtgt ctg   23
250 <210> SEQ ID NO: 7
251 <211> LENGTH: 21
252 <212> TYPE: DNA
253 <213> ORGANISM: Artificial Sequence
255 <220> FEATURE:
256 <223> OTHER INFORMATION: Primer for generation of human Fkh cDNA
258 <400> SEQUENCE: 7
259 agcctgccct tggacaagga c   21
261 <210> SEQ ID NO: 8
262 <211> LENGTH: 21
263 <212> TYPE: DNA
264 <213> ORGANISM: Artificial Sequence
266 <220> FEATURE:
267 <223> OTHER INFORMATION: Primer for generation of human Fkh cDNA
269 <400> SEQUENCE: 8
270 gcaagacagt ggaaacctca c   21

RAW SEQUENCE LISTING ERROR SUMMARY   DATE: 07/18/2002
PATENT APPLICATION: **US/09/697,340**   TIME: 13:00:20

Input Set : **A:\501d4.app**
Output Set: **N:\CRF3\07182002\I697340.raw**

<u>Please Note:</u>
Use of n and/or Xaa have been detected in the Sequence Listing. Please review the
Sequence Listing to ensure that a corresponding explanation is presented in the <220>
to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:13; N Pos. 7

VERIFICATION SUMMARY
PATENT APPLICATION: **US/09/697,340**

DATE: 07/18/2002
TIME: 13:00:20

Input Set : **A:\501d4.app**
Output Set: **N:\CRF3\07182002\I697340.raw**

L:324 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!
L:328 M:258 W: Mandatory Feature missing, <220> not found for SEQ ID#:13
L:329 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:13 after pos.:0